Early Improvement During Manual-Guided Cognitive and Dynamic Psychotherapies Predicts 16-Week Remission Status

Paul Crits-Christoph, Ph.D.
Mary Beth Connolly, Ph.D.
Robert Gallop, Ph.D.
Jacques P. Barber, Ph.D.
Xin Tu, Ph.D.
Madeline Gladis, Ph.D.
Lynne Siqueland, Ph.D.

This study examined the extent to which improvement from baseline to weeks 2, 3, and 4 on the Beck Depression Inventory and Beck Anxiety Inventory predict week 16 clinical remission for patients with major depressive disorder, generalized anxiety disorder, and/or obsessive-compulsive or avoidant personality disorders who were receiving manual-based psychotherapies. Logistic regression and receiver-operator characteristic analyses revealed relatively accurate identification of remitters and nonremitters based on improvement from baseline to sessions 2 to 4 in both original and cross-validation samples. Predictive success did not vary as a function of diagnosis, treatment type (cognitive or dynamic), or treatment status (short-term or long-term). The clinical implications of the results are discussed.

(The Journal of Psychotherapy Practice and Research 2001; 10:145–154)

Identifying patients who are likely to fail to respond to a specific form of treatment is an important challenge that has direct clinical implications. One strategy for potentially identifying failures in psychotherapy is based on comparing the pattern of improvement for individual patients with the expected rate of change for other patients with similar pre-treatment prognostic characteristics.[1] However, improvement from psychotherapy is generally not highly predictable from pre-treatment characteristics.[2]

An alternative to predicting outcome from pre-treatment characteristics is to examine early response as a predictor. If likely treatment failures can be identified early in the treatment process, alternative treatment strategies can be pursued. For example, an individual case could be reviewed so that techniques could be altered or increased clinical supervision added. If a patient has a disorder for which there are

Received November 27, 2000; revised March 17, 2001; accepted March 20, 2001. From the Department of Psychiatry, University of Pennsylvania, Room 650, 3535 Market Street, Philadelphia, PA 19104. Address correspondence to Dr. Crits-Christoph.

efficacious medications, early identification of likely therapy failures could lead to a decision to add medication to the psychotherapy treatment or to switch to medication. The identification of early failures during the treatment of major depressive disorder would be especially important given the risk for suicide that accompanies depression and the potential legal implications of not considering alternative or enhanced treatments when psychotherapy is failing. Moreover, identification of early failures has important cost implications.

Identification of potential failures in manual-guided psychotherapy, in particular, has both clinical and research implications. Knowledge of the limits of these approaches (i.e., when they are likely to fail) would be highly useful in conveying a more realistic sense of the application of manual-guided treatments in clinical practice. Although therapists may be aware of published treatment response rates for manual-based psychotherapies, therapists probably begin, and continue, therapy with most new patients by taking the view that success will occur with the implementation of an empirically supported treatment, rather than by taking a probabilistic view of success. A method indicating, for example, that at session 4 the current patient has an 80% chance of failure is more likely to raise consideration of alternative treatments.

In a research context, the early identification of likely nonresponders could be used to design trials that would allow investigators to evaluate the efficacy of alternative treatments for nonresponders without having to wait until a full course of treatment is administered. This approach would be cost-effective for research as well as reduce the amount of time that symptomatic patients have to endure an ongoing ineffective treatment before an alternative is considered.

Two studies have examined the relationship of early treatment response to final outcome for cognitive therapy (CT) of depression. Fennell and Teasdale[3] report that 8 of 17 patients treated with CT decreased 50% or more in Beck Depression Inventory[4] (BDI) scores from baseline to session 4. All 8 of these patients received a BDI score less than 10 at the termination of treatment. Beckham[5] describes the results of 23 patients in CT for depression who were classified as early responders or nonresponders at session 6 based on greater than or less than 50% improvement from baseline to session 6 on the BDI. The average BDI score for early responders at termination was 6.9, whereas nonresponders had an average BDI at termination of 27.2. Although these studies suggest a strong relationship between early improvement and final treatment response in CT for depression, both studies had small sample sizes. In addition, whether or not this relationship applies to other manual-based psychotherapies, other disorders besides depression, and equally to patients in longer-term and short-term treatment has not been investigated. A number of additional articles have recently addressed the issue of rapid early improvement in manual-based cognitive therapy for major depressive disorder, but these articles address the issue of average improvement levels, not individual differences in early response to treatment, as an indicator of eventual treatment response.[6–9]

No studies have examined whether early nonresponders to manual-guided psychotherapy might respond to an alternative treatment. However, a recent small study found that nonresponders (assessed at termination) to manual-based cognitive-behavioral and interpersonal psychotherapy for bulimia nervosa subsequently responded more to fluoxetine than to placebo.[10] Early nonresponse has been investigated in psychopharmacological studies, in which week 2 nonresponse to fluoxetine treatment of major depression was found to be associated with a 0.45 probability of nonresponse at week 8.[11]

The examination of conditional probabilities of final response or remission given early response or nonresponse, as reported in the study of fluoxetine,[11] has particular clinical appeal. Such conditional probabilities can be further enhanced by employing receiver-operator characteristic (ROC) curve analyses[12,13] that examine sensitivity (probability of a true positive response) and specificity (probability of a true nonresponse) at varying levels of the predictor variable. Most applications of ROC analysis in medicine have been confined to the context of evaluating how well a screening test predicts the presence/absence of a disorder. Our purpose in using this tool here is to supplement significant statistical findings with clinically useful guidelines for treatment decisions. These additional analyses help translate dimensionless quantities, such as effect size estimates or *P*-values, into a more interpretable form, such as the likelihood of eventual remission of symptoms based on early treatment response or nonresponse. A similar approach has been previously applied to the pharmacological treatment of obsessive-compulsive disorder,[14] although no other studies have applied ROC

analysis to psychotherapeutic treatment studies to address the relationship of early response to final remission.

The purpose of the present study was to examine early change in symptoms as a predictor of subsequent remission status during manual-guided cognitive and psychodynamic treatment of patients with elevated levels of anxiety and depressive symptoms. Specifically, we combined data from eight studies to ask the following question: to what extent does improvement in depressive and anxiety symptoms from baseline to weeks 2, 3, and 4 predict clinical remission at 4 months? To this end, we applied logistic regression to predict remission status from baseline and weekly symptom measures. The obtained logistic regression equation was then cross-validated in a second sample of patients. Derived predicted remission scores from the logistic regression analyses were employed as a "test," and sensitivity and specificity values were examined using ROC analyses. Potential clinical guidelines for decision-making are provided, based on cutoff scores obtained empirically from the ROC analyses. In examining this question of early change as a predictor of remission, we focused our attention on two instruments, the Beck Depression Inventory and the Beck Anxiety Inventory,[15] that are easy to administer on a session basis and that assess symptoms relevant to the bulk of patients presenting with anxiety, affective, and personality disorders. In addition, *a priori* definitions of clinical remission (i.e., the relative absence of symptoms) exist for these instruments.

We also examined, on an exploratory basis, the extent to which successful prediction of 4-month symptom remission varied as a function of type of treatment (cognitive or dynamic therapy), diagnosis, and eventual length of treatment (16 weeks or 1 year). We hypothesized: 1) that the relationship between early response and outcome seen for cognitive therapy in previous studies[3,5] would be stronger than that for less-structured psychodynamic therapy, in which the important insights that drive change may occur at any point in treatment; 2) that patients with chronic disorders (e.g., Axis II disorders) would demonstrate less connection between early symptom change and 16-week remission compared with those with transient (Axis I) disorders; and 3) that the connection between early response and 16-week remission was hypothesized to be less robust in long-term treatment that does not press as strongly for change early in the treatment process compared with brief therapy.

METHODS

Procedures

The present study utilized two samples. The first sample was used to derive the predictive equations and consisted of data from 8 separate single-group open trials that examined the outcome of cognitive therapy (CT) or supportive-expressive (SE) dynamic psychotherapy for one of four diagnostic groups: chronic depression (major depression, chronic subtype; or major depression plus dysthymic disorder), generalized anxiety disorder, avoidant personality disorder, and obsessive-compulsive personality disorder. Patients were recruited through a central departmental patient referral line, newspaper advertisement, personal referrals, and professional referrals. Further details of some of these pilot projects have been previously published.[16–18]

The second sample was used to cross-validate the predictive equations obtained in the first sample. The CT group from the National Institute of Mental Health Treatment of Depression Collaborative Research Program (TDCRP)[19] was used as this cross-validation sample. (Although the TDCRP study involved other treatments as well, only the CT group received weekly assessments of depressive symptoms.)

Patients

In the first sample, each patient initially participated in a 20- to 30-minute telephone screening to determine likely eligibility for one of the 8 open trials. Patients who did not meet any of the medical or psychiatric exclusion criteria were scheduled for an initial evaluation. Patients were diagnosed by using the Structured Clinical Interview for the DSM-III-R[20] and the Structured Clinical Interview for the DSM-III-R Personality Disorders.[21]

In order to limit the analyses to patients with clinically meaningful initial anxiety and depression symptoms, the current study focused on two subgroups of the total number of patients enrolled in the 8 trials. The first subgroup consisted of only the 105 patients who received a score of 16 or greater on the BDI (indicating at least moderate symptoms of depression) at baseline

and attended at least 10 treatment sessions (18 patients having dropped out prior to session 10). The second subgroup consisted of those patients ($n=79$) who received a score of 16 or greater on the BAI (indicating at least moderate symptoms of anxiety) at intake and attended at least 10 treatment sessions (18 also having dropped out prior to session 10). For analyses of BDI and BAI at specific treatment sessions, the samples are slightly lower than the respective full samples of 105 and 79 because of missing data.

A full description of the patients in the TDCRP (the validation sample for the current study) is given elsewhere.[19]

Treatments

For the derivation sample, patients with diagnoses of chronic depression or generalized anxiety received 16 weekly sessions followed by 3 monthly booster sessions of either SE psychotherapy or CT. The SE psychotherapy was based on a general SE treatment manual[22] in conjunction with diagnostically specific manuals for the treatment of generalized anxiety disorder[23] and of depression.[24] The CT therapy was conducted according to manuals for depression[25] and anxiety.[26]

The patients with avoidant or obsessive-compulsive personality disorders received 52 weekly sessions of SE or CT. SE again was based on a general SE manual, as well as unpublished manuals addressing issues specific to patients with personality disorders. CT was conducted according to a personality disorder manual.[27]

Patients in the cross-validation sample (TDCRP) received 20 sessions of manual-based CT[25] within 16 weeks.

Therapists and Supervision

For the derivation sample, there were 18 therapists who delivered SE therapy (5 with M.D.; 11 Ph.D., Psy.D., or Ed.D.; 2 M.A.) and 16 therapists (primarily graduates of a postdoctoral clinical fellowship in cognitive therapy) who delivered CT (15 with Ph.D., Psy.D., or Ed.D.; 1 M.A.). The CT and SE supervisors were experienced psychotherapists with extensive experience in training therapists in manual-based therapies. All therapists received 1 hour of individual supervision for every 2 hours of therapy provided.

Description of therapists, training, and supervision for the TDCRP (cross-validation sample) are provided elsewhere.[19]

Measures

All patients in the derivation sample received an extensive diagnostic battery as well as a self-report battery at treatment intake, at 4 and 8 months for the longer-term therapy patients, at treatment termination, and at treatment follow-up. In addition, selected self-report measures were completed at the beginning of each treatment session. The present study focused only on the month 4 assessments in order to equate for time across the short-term (16 weeks) and long-term (1 year) psychotherapy studies. However, in order to maximize sample size and generalize to a larger population, patients who dropped out of treatment (BDI: $n=8$ for CT, $n=6$ for SE; BAI: $n=3$ for CT, $n=3$ for SE) between sessions 10 and 16 were included in the sample, with their last session assessment determining their remission status.

The BDI consists of 21 items designed to assess the common symptoms of depression. The BDI was completed by all patients in the derivation sample at intake (baseline score), at the beginning of each session, and at the 4-month evaluation. Remission status on the BDI (minimal to no symptoms of depression) was defined as a 4-month score less than 10.[28]

The BAI is a 21-item self-report inventory designed to assess the common features of anxiety. The BAI was completed by all patients in the derivation sample at intake, at the beginning of each session, and at the 4-month evaluation. As on the BDI, remission status was defined on the BAI as a 4-month score less than 10.[29]

In the cross-validation sample, the BDI was administered at intake, at each treatment session, and at termination (16 weeks). Remission was defined in the same way as in the derivation sample. As in the derivation sample, only patients who had BDI scores for session 10 or thereafter were included ($n=46$ out of 59 patients who began treatment).

Statistical Analyses

Using the derivation sample, logistic regression analyses were conducted to examine the relationship between early improvement and 16-week remission status. Each logistic regression analysis had two predictors: the baseline BAI or BDI score and one of the

weekly (session 2, 3, or 4) BAI or BDI scores. Separate logistic regressions were performed for the BAI and BDI scores at weeks 2, 3, and 4 (i.e., a total of 6 logistic regression analyses). When baseline scores are included in the logistic regression, the statistical significance of the logistic regression parameter estimate for the week 2, 3, or 4 BAI or BDI measure reflects the effect of early change, corrected for initial level, on 16-week remission status.

If the logistic regression equation is to be implemented in a clinical context, two additional statistics are useful. The first is a cutoff score for the index of early improvement that would inform a clinician as to whether a given patient was likely to be a remitter or nonremitter. The second is the probability of a patient's being a remitter/nonremitter if he or she scores above or below the cutoff. To derive these statistics, the actual baseline and specified weekly BAI or BDI scores for all patients are entered into the logistic regression equation to yield "predicted remission scores." These predicted remission scores are the particular combination of baseline and specified weekly BAI or BDI scores that best predicts 16-week remission status. Conceptualizing this index of early improvement as a predictive "test" of remission status, we applied ROC analysis. Within the context of the present study, ROC curves provide plots of estimates of the conditional probability of final remission given early improvement (sensitivity) or relative nonimprovement (1 minus specificity) for each cutoff in the range of early treatment improvement. Thus, the curve provides a convenient visual display of the predictive outcomes as well as the tradeoffs between sensitivity and specificity for a given cutoff. In addition, the ROC curve has an appealing analytical property: the area under the ROC curve (AUC) also provides an indication of classification accuracy. An AUC of 0.50 indicates chance levels of prediction of the outcome criteria, and a larger AUC indicates better classification or decision rules.

In addition to evaluating the AUC, we determined an optimal cutoff for the predictive measure by simultaneously maximizing both sensitivity and specificity.[30] Sensitivity and specificity values were also calculated by using a cutoff score that increased the accuracy of identification of nonresponders (i.e., sensitivity of approximately 0.80). In addition to sensitivity and specificity, the predictive value of a positive test (PVP) and predictive value of a negative test (PVN) were calculated. The former is the percentage of subjects with a positive test (i.e., "test," as defined above, indicates likely remission) that actually do remit, while the latter is the percentage of subjects with a negative test (i.e., likely not to remit) who actually do not remit.

The logistic regression equation coefficients (and constant) calculated on the original sample were applied to data from the cross-validation sample. The two most useful clinical statistics for identifying nonremitters, specificity and PVN, were then also calculated for the cross-validation sample by using the optimal cutoff from the derivation sample that achieved a specificity of 0.80.

RESULTS

Demographic and Descriptive Characteristics of Samples

In the derivation sample, about 30% of patients received CT, and 70% participated in studies of SE psychotherapy. Sixty percent of patients received short-term psychotherapy and 40% participated in the longer-term 52-week treatments. Pooled across the 8 open trials, the sample was diagnostically heterogeneous, with substantial comorbidity. Fifty-one percent received a diagnosis of major depressive disorder (MDD), 39% generalized anxiety disorder (GAD), 34% avoidant personality disorder (AVPD), and 20% obsessive-compulsive personality disorder (OCPD). Other common diagnoses included social phobia (28% of the sample) and simple phobia (15%). In addition, 63% of patients received at least one personality disorder diagnosis.

The mean age of the derivation sample was 37 years (SD = 10, range 19 to 65). Fifty-four percent of patients were females. Eleven percent of the sample were African American, 2% were Asian, 2% were Hispanic, and 81% were white. Forty-seven percent of the sample were married or cohabiting, 37% were single, separated, or widowed, and 14% were divorced. Seventy-eight percent of patients had at least some college education. At treatment intake, the mean BDI score of those with initial BDIs of 16 or greater was 25.2 (SD = 7.3). The mean intake BAI score for the subsample with initial BAIs of 16 or greater was 26.9 (SD = 8.2). Sixty-one patients had both an intake BDI and BAI of 16 or greater, and therefore there was considerable overlap between the two samples.

Regarding outcome, 44.8% of patients met criteria for remission on the BDI at 4 months, and 60.8% met remission criteria on the BAI.

Logistic Regressions

Table 1 presents the results of the logistic regression analyses from the derivation sample. The results indicated that change on the BDI from baseline to each of weeks 2, 3, and 4 significantly predicted remission at week 16 (all $P<0.005$). Similarly, change on the BAI from baseline to each of weeks 2, 3, and 4 significantly predicted remission at week 16 (all $P<0.005$).

ROC Analyses

For illustrative purposes, one example of an ROC curve is given in Figure 1. Each point on the ROC curve is the sensitivity plotted against the specificity (actually, 1 minus specificity) for every possible cutoff score of the predictor. The "predictor" in the case of Figure 1 is a new variable ("predicted remission") created through combining baseline and week 2 BDI scores, weighting each using the coefficients from the logistic regression. As seen in Figure 1, the obtained data produces a curve that is above and to the left of the diagonal line, indicating that the "predictive test" yields better prediction than chance (as would be expected given the significant results from the logistic regression).

Table 2 gives the full descriptive results of the ROC analyses for BDI and BAI, for the equations involving baseline and each weekly score. The ROC analyses revealed that the area under the curve was in the fair to good range[31] for both BAI and BDI, and for change to weeks 2, 3, and 4. Optimal cutoffs for the predicted remission scores were determined as described in the Statistical Analyses section and are given in Table 2. With the early improvement predictive measure now a binary measure (above and below the cutoff), the sensitivity and specificity of this predictive test in relation to 16-week remission can be calculated, and these values are also shown in Table 2. As can be seen, there were only slight differences among the weeks in sensitivity and specificity values. Thus, the strength of early response as a predictor of final remission was similar for change from baseline to either session 2, 3, or 4.

Sensitivity and specificity values based on selecting a cutoff that allows for good identification of nonresponders (i.e., specificity of about 0.80) are given in Table 3. As expected, changing the cutoff to increase specificity necessarily lowers sensitivity, although this tradeoff may be acceptable if the focus is on the identification of nonresponders.

Impact of Treatment Variables on Predictions of Remission

Additional logistic regression models were included to examine the impact of treatment type (SE vs. CT), diagnosis (GAD, MDD, any Axis II), and treatment length (16 weeks vs. 1 year) in interaction with early change as predictors of 16-week remission status on the BDI and BAI. In order to limit the number of analyses, only change from baseline to session 3 was examined. Session 3 was selected because it yielded slightly better sensitivity and specificity values than sessions 2 or 4 for the BDI. In addition, only the more common diagnoses (GAD and MDD) were selected, since other diagnoses occurred at a lower frequency that would limit statistical power. Main effects for treatment type, treatment length, presence of GAD diagnosis, presence of MDD diagnosis, and presence of any Axis II diagnosis were all nonsignificant. Analyses including interaction terms revealed no evidence for differential prediction for SE vs. CT, for presence vs. absence of a GAD diagnosis, or for presence/absence of any Axis II diagnosis. A significant interaction did emerge for presence/absence of an MDD diagnosis for the BDI

TABLE 1. Logistic regression results predicting week 16 remission based on change from baseline to week 2, 3, and 4

	Week 1			Week 2			Week 3		
Measure	χ^2	**Sample Size**	***P***	χ^2	**Sample Size**	***P***	χ^2	**Sample Size**	***P***
BDI	11.1	96	0.0009	15.3	99	0.0001	10.0	98	0.0016
BAI	11.1	74	0.0009	10.5	74	0.001	8.3	73	0.004

Note: All logistic regression models include baseline measures in the regression model, although significance levels are presented only for the week 2, 3, and 4 measures, which reflect change corrected for baseline level. All chi-square tests have one degree of freedom. BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory.

($\chi^2 = 5.2$, df = 1, $P = 0.02$), but not for the BAI ($\chi^2 = 0.8$, df = 1, $P = 0.38$). Separate logistic regression analyses and ROC analyses for those patients with and without an MDD diagnosis revealed very high prediction (AUC = 0.90) for patients without a diagnosis of MDD and good prediction (AUC = 0.72) for patients with a diagnosis of MDD.

Cross-Validation of Predictive Equation

Individual patients' BDI scores at baseline and session 3 in the cross-validation sample (TDCRP) were input into the logistic regression equation that was determined on the full derivation sample using baseline and session 3 as predictors. This equation was: [predicted remission score = 1.9 + (0.03 × baseline BDI) − (0.16 × session 3 BDI)]. Using the optimal cutoff of –0.0007 obtained in the derivation sample for a specificity of 0.80, a specificity of 0.71 and a PVN of 0.86 was obtained in the cross-validation sample. A logistic regression calculated within the TDCRP sample resulted in a very similar equation: [1.5 + (0.03 × baseline BDI) − (0.13 × session 3 BDI)].

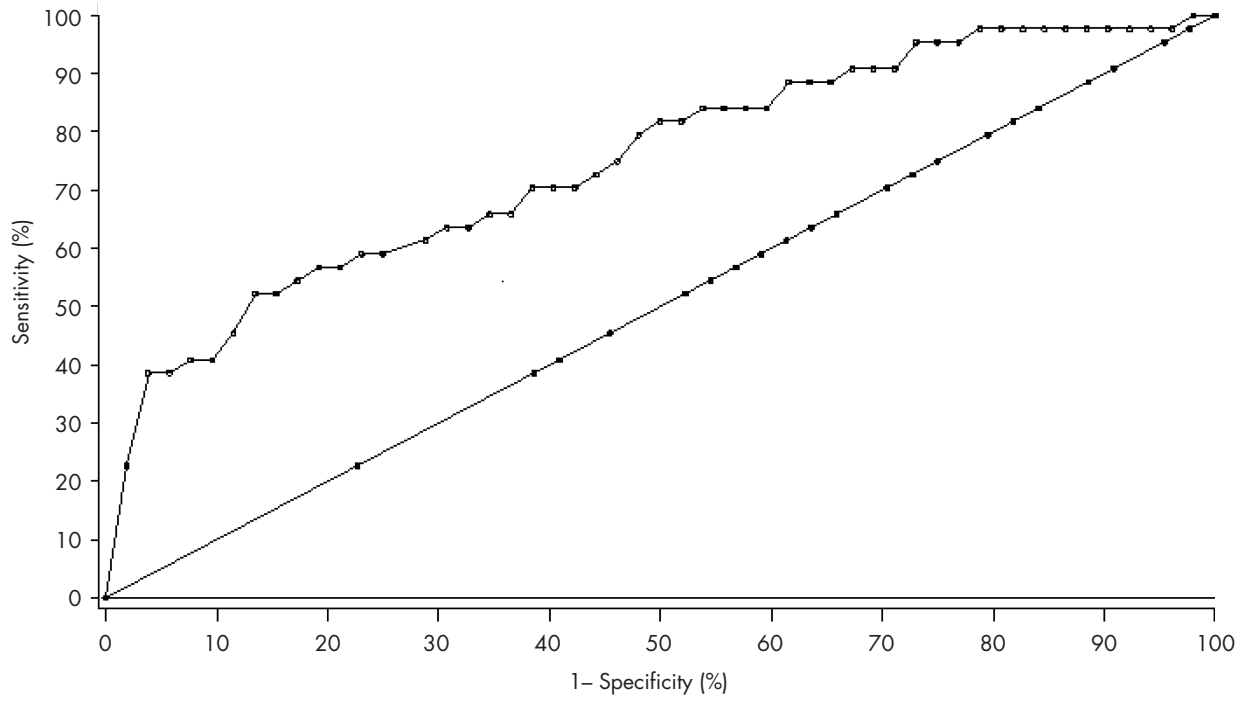
Implementation of Predictive Model

To implement the predictive model clinically, an individual patient's "predicted remission" score is calculated by using that patient's baseline and weekly scores on the BDI or BAI. For example, using the above derivation sample equation for baseline and session 3 BDI scores and the cutoff score for 0.80 specificity, a patient with a baseline BDI of 25 and a week 3 BDI of 18 would have a predicted remission score of –0.23 [1.9 + (0.03 × 25) − (0.16 × 18)]. Since the score for this patient is below the cutoff of –0.0007, the patient is predicted to not remit by week 16. The probability of not remitting is given by the predicted value negative (PVN) statistic, which is 0.76 (see PVN for week 3 BDI in Table 3).

DISCUSSION

Early improvement from baseline to weeks 2 through 4 was highly associated with clinical remission in cognitive and psychodynamic treatment of MDD, GAD, AVPD, and OCPD. Using an optimal cutoff score that

FIGURE 1. Receiver-operator characteristic curve for early improvement in Beck Depression Inventory score (from baseline to session 2) predicting remission of symptoms at week 16. The data produce a curve that is above and to the left of the diagonal line, indicating that the "predictive test" yields better prediction than chance.



attempts to maximize both sensitivity and specificity, the statistical model identified potential remitters on the BDI with a relatively high degree of accuracy (about 80% accuracy) as early as week 3. With the BAI, the optimal cutoff identified nonremitters the best (83% accuracy at week 4). The equation for the BDI examining change to session 3 received strong cross-validation in a second sample. Although it has always been known that some patients do well and some do not do well in psychotherapy, the current data are the first to indicate that across a range of patient diagnoses, treatment types, and treatment lengths, eventual remission/nonremission of symptoms is highly predictable from the early pattern of treatment response.

The mechanism for the strong relationship between early improvement and final remission status is not clear. This relationship is not necessarily linked to a specific psychotherapy "technique" effect. Studies of the relationship of early response to pharmacological treatment outcome with major depression have suggested that the effect is more related to the nonspecific treatment factors than to a specific drug effect.[32,33] Thus, it may be largely "placebo" responders who evidence marked improvement in the first few treatment sessions and then continue to improve to the point of remission by week 16. Another possibility is that early improvement reflects the natural course of the disorder, rather than anything to do with treatment per se. Disorders that have a more transient nature might show rapid early improvement (continuing through to termination), while chronic disorders might not. Arguing against this explanation is the fact that the natural course of untreated GAD, MDD, AVPD, and OCPD (the primary disorders in the sample) is not likely to be rapid improvement within two weeks. Moreover, the predictions did not vary by presence/absence of Axis II disorders, and Axis II disorders are by definition chronic.

The clinical implications of these results relate to the potential for modifying treatment for those patients who early on are highly predicted to be nonresponders. If a patient is likely to be a nonresponder, alternative

TABLE 2. ROC analysis with sensitivity and specificity based on optimal cutoff

Measure	Area Under the Curve ± SE	Optimal Cutoff				
		Cutoff	Sensitivity	Specificity	PVP	PVN
BDI						
Week 2	0.75 ± 0.07	–0.04	0.66	0.65	0.62	0.69
Week 3	0.79 ± 0.06	–0.17	0.73	0.75	0.70	0.77
Week 4	0.73 ± 0.07	–0.10	0.68	0.69	0.64	0.73
BAI						
Week 2	0.76 ± 0.08	0.46	0.71	0.69	0.78	0.61
Week 3	0.76 ± 0.08	0.52	0.73	0.73	0.80	0.65
Week 4	0.77 ± 0.08	0.54	0.75	0.79	0.85	0.68

Note: Results are given for the linear combination of baseline and week 2, 3, or 4 BDI or BAI score as a predictor of clinical remission (on BDI or BAI, respectively) at week 16. Optimal cutpoint is determined by a method that simultaneously maximizes both sensitivity and specificity.[30] ROC = receiver-operator characteristic; BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; PVP = predictive value of a positive test; PVN = predictive value of a negative test.

TABLE 3. Sensitivity and specificity based on cutoff that produces better identification of nonremitters

Measure	Sensitivity	Specificity	PVP	PVN
BDI				
Week 2	0.57	0.81	0.71	0.69
Week 3	0.68	0.80	0.73	0.76
Week 4	0.57	0.80	0.69	0.69
BAI				
Week 2	0.62	0.79	0.82	0.58
Week 3	0.61	0.80	0.82	0.59
Week 4	0.66	0.83	0.85	0.62

Note: Results are given for the linear combination of baseline and week 2, 3, or 4 BDI or BAI score as a predictor of clinical remission (on BDI or BAI, respectively) at week 16. Cutoff was chosen to yield specificity of about 0.80. BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory PVP = predictive value of a positive test; PVN = predictive value of a negative test.

treatments (other psychosocial treatments or medications) can be considered, either instead of the current treatment or as an augmentation strategy (combined treatment). If a clinician has data that suggest with a high degree of probability that a given patient will not respond to the current treatment, it may be worth considering other options that may yield greater benefit. However, no empirical research exists in regard to the efficacy/effectiveness of alternative treatments for nonresponders to psychotherapy. It may be that such nonresponders will fail in all treatments, and clinicians may be reluctant to disrupt the current therapeutic relationship without hope that a different treatment will be better. Augmentation strategies may be most practical here, allowing the current therapeutic relationship to continue while adding another treatment (e.g., medication) that has some possibility of improving response.

Several limitations of these analyses should be mentioned. Although the sensitivity and specificity values obtained in this study were generally good, there was still room for improvement, and some clinicians may prefer values at or above 90% before relying on a statistical model to guide clinical decision-making. However, the predictive accuracy of a "test" needs to be evaluated relative to the reproducibility of the "gold standard" outcome. Within depressed samples, the test–retest reliability of the BDI has ranged from 0.65[34] to 0.86.[35] Short-term test–retest of the BAI has been reported to be 0.75.[15] Considering the reliability constraints of these instruments, there is only minimal room for improving the sensitivity and specificity values of the predictive equation. However, to the extent that prediction success can be increased, future research could include additional (e.g., pre-treatment) predictors of outcome such as those identified by Howard et al.[1] in order to enhance the identification of responders and nonresponders. Additional process variables, such as the therapeutic alliance, could also be considered as candidates for evaluating treatment progress in order to predict final remission.[36] However, it is likely that there is a limit on the predictive power of pre-treatment, early process, and early improvement measures in predicting final outcome. Recent research, for example, has shown that "sudden gains" occur throughout the course of CT for depression.[37] The existence of such sudden gains beyond the early phase of treatment would tend to decrease the relationship of early improvement to final remission status.

Another limitation is that the analyses focused solely on symptom-based outcome measures. In making the decision about whether to change or augment treatment, it may be important to consider other types of outcomes (i.e., functioning and quality of life). However, functioning and quality of life indices do not change as markedly within a short duration of psychotherapy,[38,39] and most managed care agencies allow for only a relatively short duration of treatment (e.g., 10 to 20 sessions). Furthermore, some managed care agencies define the need for psychotherapy in terms of "medical necessity," which translates into a focus on diagnosis or symptoms related to diagnosis.

A further limitation is that the number of weeks until remission was not fully controlled in regard to the assessment of remission status. A small number of patients in the sample had their endpoint (remission) evaluation between 10 and 16 weeks. Kraemer[13] has explained that misleading ROC results can be found if prevalence rates vary over time. In the current context, if remission rates varied significantly between sessions 10 and 16, this confound of time would be a problem. However, within the derivation sample most of the change in BDI (85%) and BAI (86%) had occurred by session 10, so it appears this issue would have little effect on the results.

Although derivation sample sizes were reasonably large ($n = 105$ for BDI analyses; $n = 79$ for BAI analyses), the study may have had limited statistical power for the detection of subgroup differences (interactions). Power would be more of an issue, however, if the overall results were weak or negative, masking underlying subgroup differences. Our results showed strong prediction for the sample as a whole. When subgroup differences were detected (i.e., for presence/absence of MDD), they revealed very high prediction for one subgroup versus moderately high for the other. However, it is possible that for diagnoses not included or infrequent within the sample, a different pattern of prediction would emerge. Furthermore, there may be other patient characteristics besides diagnosis that influence the degree of relationship between early response and final remission of symptoms.

The current results may not generalize to other forms of manual or non–manual-based psychotherapies. Further research is necessary to examine the generalizability of these findings to other forms of psychotherapy, outcome assessments, and patient populations.

REFERENCES

1. Howard KI, Moras K, Brill PL, et al: Evaluation of psychotherapy: efficacy, effectiveness, and patient progress. Am Psychol 1996; 51:1059–1064
2. Luborsky L, Crits-Christoph P, Mintz J, et al: Who Will Benefit From Psychotherapy? New York, Basic Books, 1988
3. Fennell MJV, Teasdale JD: Cognitive therapy for depression: individual differences and the process of change. Cogn Ther Res 1987; 11:253–271
4. Beck AT, Steer RA, Garbin MG: Psychometric properties of the Beck Depression Inventory: twenty-five years later. Clin Psychol Rev 1988; 8:77–100
5. Beckham EE: Improvement after evaluation in psychotherapy of depression: evidence of a placebo effect. J Clin Psychiatry 1989; 45:945–950
6. Ilardi SS, Craighead WE: The role of nonspecific factors in cognitive-behavior therapy for depression. Clinical Psychology: Science and Practice 1994; 1:138–156
7. Tang TZ, DeRubeis RJ: Reconsidering rapid early response in cognitive behavioral therapy for depression. Clinical Psychology: Science and Practice 1999; 6:283–288
8. Hollon SD: Rapid early response in cognitive behavior therapy: a commentary. Clinical Psychology: Science and Practice 1999; 6:305–309
9. Ilardi SS, Craighead WE: Rapid early response, cognitive modification, and nonspecific factors in cognitive behavior therapy for depression: a reply to Tang and DeRubeis. Clinical Psychology: Science and Practice 1999; 6:295–299
10. Walsh BT, Agras WS, Devlin MJ, et al: Fluoxetine for bulimia nervosa following poor response to psychotherapy. Am J Psychiatry 2000; 157:1332–1334
11. Nierenberg AA, McLean NE, Alpert JE, et al: Early non-response to fluoxetine as a predictor of poor 8-week outcome. Am J Psychiatry 1995; 152:1500–1503
12. Green DM, Swets JA: Signal Detection Theory and Psychophysics. New York, Wiley, 1996
13. Kraemer HC: Evaluating Medical Tests: Objective and Quantitative Guidelines. Newbury Park, CA, Sage, 1992
14. Ackerman DL, Greenland S, Bystritsky A: Use of receiver-operator characteristic (ROC) curve analysis to evaluate predictors of response to clomipramine therapy. Psychopharmacol Bull 1996; 32:157–165
15. Beck AT, Epstein, N, Brown G, et al: An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988; 56:893–897
16. Luborsky L, Diguer L, Cacciola J, et al: Factors in outcomes of short-term dynamic psychotherapy for chronic vs. nonchronic major depression. J Psychother Pract Res 1996; 5:152–159
17. Crits-Christoph P, Connolly MB, Azarian K, et al: An open trial of brief supportive-expressive psychotherapy in the treatment of generalized anxiety disorder. Psychotherapy 1996; 33:418–430
18. Barber JP, Morse JQ, Krakauer I, et al: Change in obsessive-compulsive and avoidant personality disorders following time-limited supportive-expressive therapy. Psychotherapy 1997; 34:133–143
19. Elkin I, Shea MT, Watkins JT, et al: National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. Arch Gen Psychiatry 1989; 46:971–982
20. Spitzer RL, Williams JBW, Gibbon M, et al: Structured Clinical Interview for DSM-III-R-patient edition (with psychotic screen): SCID-P (with psychotic screen) (version 1.0). Washington, DC, American Psychiatric Press, 1990
21. Spitzer RL, Williams JBW, Gibbon M, et al: Structured Clinical Interview for DSM-III-R-patient edition (with psychotic screen): personality disorders (SCID-II, Version 1.0). Washington, DC, American Psychiatric Press, 1990
22. Luborsky L: Principles of Psychoanalytic Psychotherapy: A Manual for Supportive-Expressive Treatment. New York, Basic Books, 1984
23. Crits-Christoph P, Crits-Christoph K, Wolf-Palacio D, et al: Brief supportive-expressive psychodynamic therapy for generalized anxiety disorder, in Dynamic Therapies for Psychiatric Disorders (Axis I), edited by Barber JP, Crits-Christoph P. New York, Basic Books, 1995, pp 43–83
24. Luborsky L, Mark D, Hole AV, et al: Supportive-expressive dynamic psychotherapy of depression: a time-limited version, in Dynamic Therapies for Psychiatric Disorders (Axis I), edited by Barber JP, Crits-Christoph P. New York, Basic Books, 1995, pp 13–42
25. Beck AT, Rush AJ, Shaw BF, et al: Cognitive Therapy of Depression. New York, Guilford, 1979
26. Beck AT, Emery G: Anxiety Disorders and Phobias: A Cognitive Perspective. New York, Basic Books, 1985
27. Beck AT, Freeman AM: Cognitive Therapy of Personality Disorders. New York, Guilford, 1990
28. Beck AT, Steer RA: Beck Depression Inventory manual. San Antonio, TX, Psychological Corporation, 1987
29. Beck AT, Steer RA: Beck Anxiety Inventory manual. San Antonio, TX, Psychological Corporation, 1987
30. Riddle DL, Stratford PW: Interpreting validity indexes for diagnostic tests: an illustration using the Berg Balance Test. Physical Therapy 1999; 79:939–950
31. Metz CE: Basic principles of ROC analysis. J Semin Nucl Med 1978; 8:283–298
32. Priebe S, Brekker M: Initial response to active drug and placebo predicts outcome of antidepressant treatment. Eur Psychiatry 1997; 12:28–33
33. Quitkin FM, McGrath PJ, Stewart JW, et al: Can the effects of antidepressants be observed in the first two weeks of treatment? Neuropsychopharmacology 1996; 15:390–394
34. Bailey J, Coppen A: A comparison between the Hamilton Rating Scale and the Beck Inventory in the measurement of depression. Br J Psychiatry 1976; 128:486–489
35. Gallagher D, Nies G, Thompson LW: Reliability of the Beck Depression Inventory with older adults. J Consult Clin Psychol 1982; 50:152–153
36. Samstag LW, Batchelder ST, Muran JC, et al: Early identification of treatment failures in short-term psychotherapy: an assessment of therapeutic alliance and interpersonal behavior. J Psychother Pract Res 1998; 7:126–143
37. Tang TZ, DeRubeis RJ: Sudden gains and critical sessions in cognitive-behavioral therapy for depression. J Consult Clin Psychol 1999; 67:894–904
38. Kopta SM, Howard KI, Lowry JL, et al: Patterns of symptomatic recovery in psychotherapy. J Consult Clin Psychol 1994; 62:1009–1016
39. Howard KI, Lueger RJ, Maling MS, et al: A phase model of psychotherapy outcome: causal mediation of change. J Consult Clin Psychol 1993; 61:678–685